CLAIMS

What is claimed is:

1. A method for directing a network client requesting access to content to one of a plurality of content servers that can provide said content, comprising:

directing a network client to a said one of said content servers based on one or more cost measurements indicative of operational characteristics of the network.

2. A method as recited in claim 1, further comprising:

obtaining a new cost measurement when said network client accesses said content server; and

using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

3. A method as recited in claim 1, wherein said content servers are associated with a network server having an identity, and wherein said network client requests content from said network server, and further comprising:

mapping the identity of the network server to said content servers.

4. A method as recited in claim 1, further comprising measuring network performance between said network client and a said one of said content servers.

5. A method as recited in claim 1, wherein an attribute of network performance comprises network latency.

6. A method as recited in claim 5, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by one of said content servers.

7. A method as recited in claim 4, further comprising measuring network performance between said network client and another of said content servers.

8. A method as recited in claim 1, further comprising determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

9. A method as recited in claim 8, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

10. A method as recited in claim 1, further comprising inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said content server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

11. A method as recited in claim 1, further comprising:

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

12. A method for directing a network client requesting access to content from a network server to one of a plurality of content servers that can provide said content, each said content server having an address, said network server having an identity, said method comprising:

(a) identifying a network server associated with content requested by said network client;

(b) identifying a said one of said content servers based on said identity of said network server and one or more cost measurements indicative of operational characteristics of the network; and

(c) providing the network client with the address of said content server identified in step (b).

13. A method as recited in claim 12, further comprising:

(d) obtaining a new cost measurement when said network client accesses said content server; and

(e) using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

14. A method as recited in claim 12, further comprising measuring network performance between said network client and a said one of said content servers.

15. A method as recited in claim 12, wherein an attribute of network performance comprises network latency.

16. A method as recited in claim 15, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by said one of said content servers.

17. A method as recited in claim 14, further comprising measuring network performance between said network client and another of said content servers.

18. A method as recited in claim 12, determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

19. A method as recited in claim 18, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

20. A method as recited in claim 12, further comprising inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said content server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

R1.124

~~23.~~ 21. A method as recited in claim 12, further comprising:

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

22. A method for directing a network client requesting access to content from a network server to one of a plurality of content servers providing said content for said network server, each said content server having an address, said network server having an identity, said method comprising returning the address of a content server that may provide said content the identity of the network server based on one or more cost measurements indicative of operational characteristics of the network, obtaining a new cost measurement when said network client accesses said content server, and using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

23. A method as recited in claim 22, further comprising measuring network performance between said network client and a said one of said content server.

24. A method as recited in claim 22, wherein an attribute of network performance comprises network latency.

25. A method as recited in claim 24, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by said one of said content servers.

26. A method as recited in claim 23, further comprising measuring network performance between said network client and another of said content servers.

27. A method as recited in claim 22, further comprising determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

28. A method as recited in claim 27, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

29. A method as recited in claim 22, further comprising inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said convent server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

30. A method as recited in claim 22, further comprising:

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

31. A method for determining the physical location of a network client comprising:

(a) measuring the time that it takes for data to move from a plurality of network server locations to a network client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

32. A method for inferring operational characteristics associated with a plurality of network clients to an inferable network client, comprising:

(a) measuring network latency between a network server and a plurality of network clients;

(b) determining physical distances between said network clients and an inferable network client; and

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said network server and said inferable network client.

33. A system for directing a network client requesting access to content to one of a plurality of content servers that can provide said content, comprising:

a programmed data processor; and

programming associated with said programmed data processor for directing a network client to a said one of said content servers based on one or more cost measurements indicative of operational characteristics of the network.

34. A system as recited in claim 33, further comprising programming associated with said programmed data processor for:

obtaining a new cost measurement when said network client accesses said content server; and

using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

35. A system as recited in claim 33, wherein said content servers are associated with a network server having an identity, and wherein said network client requests content from said network server, and further comprising:

programming associated with said programmed data processor mapping the identity of the network server to said content servers.

36. A system as recited in claim 33, further comprising programming associated with said programmed data processor for measuring network performance between said network client and a said one of said content servers.

37. A system as recited in claim 33, wherein an attribute of network performance comprises network latency.

38. A system as recited in claim 37, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by one of said content servers.

39. A system as recited in claim 36, further comprising programming associated with said programmed data processor for measuring network performance between said network client and another of said content servers.

40. A system as recited in claim 33, further comprising programming associated with said programmed data processor for determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

41. A system as recited in claim 40, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

42. A system as recited in claim 33, further comprising programming associated with said programmed data processor for inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said content server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

43. A system as recited in claim 33, further comprising programming associated with said programmed data processor for

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

44. A system for directing a network client requesting access to content from a network server to one of a plurality of content servers that can provide said content, each said content server having an address, said network server having an identity, said method comprising:

(a) a programmed data processor; and

(b) programming associated with said programmed data processor for

(i) identifying a network server associated with content requested by said network client;

(ii) identifying a said one of said content servers based on said identity of said network server and one or more cost measurements indicative of operational characteristics of the network; and

(iii) providing the network client with the address of said content server identified in step (ii).

45. A system as recited in claim 44, further comprising programming associated with said programmed data processor for:

obtaining a new cost measurement when said network client accesses said content server; and

using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

46. A system as recited in claim 44, further comprising programming associated with said programmed data processor for measuring network performance between said network client and a said one of said content servers.

47. A system as recited in claim 46, wherein an attribute of network performance comprises network latency.

48. A system as recited in claim 47, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by said one of said content servers.

49. A system as recited in claim 46, further comprising programming associated with said programmed data processor for measuring network performance between said network client and another of said content servers.

50. A system as recited in claim 44, further comprising programming associated with said programmed data processor for determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

51. A system as recited in claim 50, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

52. A system as recited in claim 44, further comprising programming associated with said programmed data processor for inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said convent server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

53. A system as recited in claim 44, further comprising programming associated with said programmed data processor for:

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

54. A system for directing a network client requesting access to content from a network server to one of a plurality of content servers providing said content for said network server, each said content server having an address, said network server having an identity, said system comprising a programmed data processor and programming associated with said programmed data processor for returning the address of a content server that may provide said content the identity of the network server one or more cost measurements indicative of operational characteristics of the network, obtaining a new cost measurement when said network client accesses said content server, and using said new cost measurement as an indicator of operational characteristics of the network in connection with subsequent requests for access to said content that can be provided by said content server.

55. A system as recited in claim 54, further comprising programming associated with said programmed data processor for measuring network performance between said network client and a said one of said content servers.

56. A system as recited in claim 55, wherein an attribute of network performance comprises network latency.

57. A system as recited in claim 56, wherein network latency is measured passively by determining the time between a syn ack message sent by said network client and an ack message sent by said one of said content servers.

58. A system as recited in claim 55, further comprising programming associated with said programmed data processor for measuring network performance between said network client and another of said content servers.

59. A system as recited in claim 54, further comprising programming associated with said programmed data processor for determining the location of said network client by circular intersection and inferring network performance associated with accessing said content server to said network client based on a performance measurement to said content server in connection with another physically proximate network client.

60. A method as recited in claim 59, wherein said circular intersection comprises:

(a) measuring the time that it takes for data to move from a plurality of network server locations to said client;

(b) converting said times to distance equivalents;

(c) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(d) determining the physical location of said network client from the intersection of said circles.

61. A system as recited in claim 54, further comprising programming associated with said programmed data processor for inferring network performance of serving said network client from said content server by determining a weighted average of network performance between said convent server and other network clients based on physical proximity of said other network clients to said network client and performance of said content server for said other network clients.

62. A system as recited in claim 54, further comprising programming associated with said programmed data processor for:

(a) measuring network latency between a content server and a plurality of other network clients;

(b) determining physical distances between said other network clients and said network client;

(c) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said content server and said network client; and

(d) inferring operational characteristics associated with a plurality of network clients to said network client using said weighted average.

63. A system for determining the physical location of a network client comprising:

(a) a programmed data processor; and

(b) programming associated with said programmed data processor for

(i) measuring the time that it takes for data to move from a plurality of network server locations to a network client;

(ii) converting said times to distance equivalents;

(iii) forming a plurality of intersecting circles using said distance equivalents as the radius of circles with said network server locations as the center; and

(iv) determining the physical location of said network client from the intersection of said circles.

64. A system for inferring operational characteristics associated with a plurality of network clients to an inferable network client, comprising:

(a) a programmed data processor; and

(b) programming associated with said programmed data processor for

(i) measuring network latency between a network server and a plurality of network clients;

(ii) determining physical distances between said network clients and an inferable network client; and

(iii) computing a weighted average of said latency measurements as a function of said distances, wherein said weighed average comprises an estimate of the latency between said network server and said inferable network client.